

0590 #2
10/11 OIPE

RAW SEQUENCE LISTING

DATE: 06/08/2001

PATENT APPLICATION: US/09/811,118

TIME: 14:15:37

Input Set : C:\PAOLA\09811118.txt

Output Set: N:\CRF3\06082001\I811118.raw

SEQUENCE LISTING

ENTERED

5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: HILLMAN, JENNIFER L.

8 CORLEY, NEIL C.

9 PATTERSON, CHANDRA

11 (ii) TITLE OF INVENTION: HUMAN GLUTATHIONE PEROXIDASE-6

13 (iii) NUMBER OF SEQUENCES: 3

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

17 (B) STREET: 3174 Porter Drive

18 (C) CITY: Palo Alto

19 (D) STATE: CA

20 (E) COUNTRY: USA

21 (F) ZIP: 94304

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Diskette

25 (B) COMPUTER: IBM Compatible

26 (C) OPERATING SYSTEM: Windows

27 (D) SOFTWARE: FastSEQ for Windows Version 2.0b

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/811,118

C--> 31 (B) FILING DATE: 16-Mar-2001

32 (C) CLASSIFICATION:

35 (vii) PRIOR APPLICATION DATA:

36 (A) APPLICATION NUMBER: 09/088,549

37 (B) FILING DATE:

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: Cerrone, Michael C

42 (B) REGISTRATION NUMBER: 39,132

43 (C) REFERENCE/DOCKET NUMBER: PF-0530 US

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: 650-855-0555

47 (B) TELEFAX: 650-855-0572

48 (C) TELEX:

51 (2) INFORMATION FOR SEQ ID NO: 1:

53 (i) SEQUENCE CHARACTERISTICS:

54 (A) LENGTH: 187 amino acids

55 (B) TYPE: amino acid

56 (C) STRANDEDNESS: single

57 (D) TOPOLOGY: linear

59 (vii) IMMEDIATE SOURCE:

60 (A) LIBRARY: PROSNOT20

61 (B) CLONE: 1817518

63 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

65 Met Val Ala Ala Thr Val Ala Ala Ala Trp Leu Leu Trp Ala Ala

66 1 5 10 15

67 Ala Cys Ala Gln Gln Glu Gln Asp Phe Tyr Asp Phe Lys Ala Val Asn

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/811,118

DATE: 06/08/2001

TIME: 14:15:37

Input Set : C:\PAOLA\09811118.txt

Output Set: N:\CRF3\06082001\I811118.raw

```

68          20          25          30
69 Ile Arg Gly Lys Leu Val Ser Leu Glu Lys Tyr Arg Gly Ser Val Ser
70          35          40          45
71 Leu Val Val Asn Val Ala Ser Glu Cys Gly Phe Thr Asp Gln His Tyr
72          50          55          60
73 Arg Ala Leu Gln Gln Leu Gln Arg Asp Leu Gly Pro His His Phe Asn
74          65          70          75          80
75 Val Leu Ala Phe Pro Cys Asn Gln Phe Gly Gln Gln Glu Pro Asp Ser
76          85          90          95
77 Asn Lys Glu Ile Glu Ser Phe Ala Cys Arg Thr Tyr Ser Val Ser Phe
78          100          105          110
79 Pro Met Phe Ser Lys Ile Ala Val Thr Gly Thr Gly Ala His Pro Ala
80          115          120          125
81 Phe Lys Tyr Leu Ala Gln Thr Ser Gly Lys Glu Pro Thr Trp Asn Phe
82          130          135          140
83 Trp Lys Tyr Leu Val Ala Pro Asp Gly Lys Val Val Gly Ala Trp Asp
84          145          150          155          160
85 Pro Thr Val Ser Val Glu Glu Val Arg Leu Gln Ile Thr Ala Leu Val
86          165          170          175
87 Arg Lys Leu Ile Leu Leu Lys Arg Glu Asp Leu
88          180          185

```

90 (2) INFORMATION FOR SEQ ID NO: 2:

92 (i) SEQUENCE CHARACTERISTICS:

93 (A) LENGTH: 1072 base pairs

94 (B) TYPE: nucleic acid

95 (C) STRANDEDNESS: single

96 (D) TOPOLOGY: linear

98 (vii) IMMEDIATE SOURCE:

99 (A) LIBRARY: PROSNOT20

100 (B) CLONE: 1817518

102 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

104 GACGCCGCCA CCTCCGGAAC AAGCCATGGT GGC GCGGACG GTGGCAGCGG CGTGGCTGCT 60
105 CCTGTGGGCT GCGGCCTGCG CGCAGCAGGA GCAGGACTTC TACGACTTCA AGGCGGTCAA 120
106 CATCCGGGGC AAAGTGGTGT CGCTGGAGAA GTACCGCGGA TCGGTGTCCC TGGTGGTGAA 180
107 TGTGGCCAGC GAGTGGCGCT TCACAGACCA GCACTACCGA GCCCTGCAGC AGCTGCAGCG 240
108 AGACCTGGGC CCCCACTACT TCAACGTGCT CGCCTTCCCC TGCAACCAGT TTGGCCAACA 300
109 GGAGCCTGAC AGCAACAAGG AGATTGAGAG CTTTGCCTGC CGCACCTACA GTGTCTCATT 360
110 CCCCATGTTT AGCAAGATTG CAGTCACCGG TACTGGTGCC CATCCTGCCT TCAAGTACCT 420
111 GGCCAGACT TCTGGGAAGG AGCCACCTG GAACTTCTGG AAGTACCTAG TAGCCCCAGA 480
112 TGGAAAGGTG GTAGGGGCTT GGGACCCAAC TGTGTAGTG GAGGAGGTCA GACTCCAGAT 540
113 CACAGCGCTC GTGAGGAAGC TCATCCTACT GAAGCGAGAA GACTTATAAC CACCGCGTCT 600
114 CCTCCTCCAC CACCTCATCC CGCCACCTG TGTGGGGCTG ACCAATGCAA ACTCAAATGG 660
115 TGCTTCAAAG GGAGAGACCC ACTGACTCTC CTTCTTTTAC TCTTATGCCA TTGGTCCCAT 720
116 CATTCTTG TG GGGGAAAAAT TCTAGTATTT TGATTATTTG AATCTTACAG CAACAAATAG 780
117 GAACTCTGG CCAATGAGAG CTCTTGACCA GTGAATCACC AGCCGATACG AACGTCTTGC 840
118 CAACAAAAAT GTGTGGCAAA TAGAAGTATA TCAAGCAATA ATCTCCCACC CAAGGCTTCT 900
119 GTAACTGGG ACCAATGATT ACCTCATAGG GCTGTTGTGA GGATTAGGAT GAAATACCTG 960
120 TGAAAGTGCC TAGGCAGTGC CAGCCAAATA GGAGGCATTC AATGAACATT TTTTGCATAT 1020
121 AAACCAAAAA ATAAGTTGTT ATCAATAAAA ACTTGCATCC AACATGAATT TC 1072

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/811,118

DATE: 06/08/2001

TIME: 14:15:37

Input Set : C:\PAOLA\09811118.txt

Output Set: N:\CRF3\06082001\I811118.raw

123 (2) INFORMATION FOR SEQ ID NO: 3:

125 (i) SEQUENCE CHARACTERISTICS:

126 (A) LENGTH: 170 amino acids

127 (B) TYPE: amino acid

128 (C) STRANDEDNESS: single

129 (D) TOPOLOGY: linear

131 (vii) IMMEDIATE SOURCE:

132 (A) LIBRARY: GENBANK

133 (B) CLONE: 544432

135 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

137 Met Cys Ala Ser Arg Asp Asp Trp Arg Cys Ala Arg Ser Met His Glu
138 1 5 10 15
139 Phe Ser Ala Lys Asp Ile Asp Gly His Met Val Asn Leu Asp Lys Tyr
140 20 25 30
141 Arg Gly Phe Val Cys Ile Val Thr Asn Val Ala Ser Gln Cys Gly Lys
142 35 40 45
143 Thr Glu Val Asn Tyr Thr Gln Leu Val Asp Leu His Ala Arg Tyr Ala
144 50 55 60
145 Glu Cys Gly Leu Arg Ile Leu Ala Phe Pro Cys Asn Gln Phe Gly Lys
146 65 70 75 80
147 Gln Glu Pro Gly Ser Asn Glu Glu Ile Lys Glu Phe Ala Ala Gly Tyr
148 85 90 95
149 Asn Val Lys Phe Asp Met Phe Ser Lys Ile Cys Val Asn Gly Asp Asp
150 100 105 110
151 Ala His Pro Leu Trp Lys Trp Met Lys Ile Gln Pro Lys Gly Lys Gly
152 115 120 125
153 Ile Leu Gly Asn Ala Ile Lys Trp Asn Phe Thr Lys Phe Leu Ile Asp
154 130 135 140
155 Lys Asn Gly Cys Val Val Lys Arg Tyr Gly Pro Met Glu Glu Pro Leu
156 145 150 155 160
157 Val Ile Glu Lys Asp Leu Pro His Tyr Phe
158 165 170

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/811,118

DATE: 06/08/2001

TIME: 14:15:38

Input Set : C:\PAOLA\098111118.txt

Output Set: N:\CRF3\06082001\I8111118.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]